



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>6</sup>:</b> C12N 15/54, 15/79, 9/12, 15/10, 1/21, 1/19, C12Q 1/48, 1/68, G01N 33/573, A61K 38/17, 31/70, 31/55, 31/275, 31/35, 47/48	<b>A1</b>	<b>(11) International Publication Number:</b> WO 95/02057 <b>(43) International Publication Date:</b> 19 January 1995 (19.01.95)
<b>(21) International Application Number:</b> PCT/GB94/01479 <b>(22) International Filing Date:</b> 8 July 1994 (08.07.94)  <b>(30) Priority Data:</b> 9314233.9                      9 July 1993 (09.07.93)                      GB 9404817.0                      11 March 1994 (11.03.94)                      GB  <b>(71) Applicants (for all designated States except US):</b> THE INSTITUTE OF CANCER RESEARCH [GB/GB]; 17A Onslow Gardens, London SW7 3AL (GB). THE WELLCOME FOUNDATION LIMITED [GB/GB]; Unicorn House, 160 Euston Road, London NW1 2BP (GB).  <b>(72) Inventors; and</b> <b>(75) Inventors/Applicants (for US only):</b> GUSTERSON, Barry, Austin [GB/GB]; Institute of Cancer Research: Royal Cancer Hospital, Haddow Laboratories, 15 Cotswold Road, Sutton, Surrey SM2 5NG (GB). CROMPTON, Mark, Roger [GB/GB]; Institute of Cancer Research: Royal Cancer Hospital, Haddow Laboratories, 15 Cotswold Road, Sutton, Surrey SM2 5NG (GB). MITCHELL, Philip, John [GB/GB]; Institute of Cancer Research: Royal Cancer Hospital, Haddow Laboratories, 15 Cotswold Road, Sutton, Surrey SM2 5NG		<b>(GB).</b> BARKER, Karen, Tracey [GB/GB]; Institute of Cancer Research: Royal Cancer Hospital, Haddow Laboratories, 15 Cotswold Road, Sutton, Surrey SM2 5NG (GB). KAMALATI, Tahereh [GB/GB]; Institute of Cancer Research: Royal Cancer Hospital, Haddow Laboratories, 15 Cotswold Road, Sutton, Surrey SM2 5NG (GB). PAGE, Martin, John [GB/GB]; The Wellcome Research Laboratories, Langley Court, South Eden Park Road, Beckenham, Kent BR3 3BS (GB). SPENCE, Paul [GB/GB]; The Wellcome Research Laboratories, Langley Court, South Eden Park Road, Beckenham, Kent BR3 3BS (GB).  <b>(74) Agent:</b> MARCHANT, James, Ian; Elkington and Fife, Prospect House, 8 Pembroke Road, Sevenoaks, Kent TN13 1XR (GB).  <b>(81) Designated States:</b> AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, ES, FI, GB, GE, HU, JP, KE, KG, KP, KR, KZ, LK, LT, LU, LV, MD, MG, MN, MW, NL, NO, NZ, PL, PT, RO, RU, SD, SE, SI, SK, TJ, TT, UA, US, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).  <b>Published</b> <i>With international search report.</i> <i>Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.</i>
<b>(54) Title:</b> PROTEIN TYROSINE KINASE AND LIGANDS THEREOF  <b>(57) Abstract</b>  The invention relates to a novel tyrosine kinase designated <u>brk</u> and the use of this tyrosine kinase in the development of diagnostic and therapeutic approaches to cancer. The novel tyrosine kinase was isolated from a human metastatic breast tumour using a PCR based differential screening approach and has the deduced amino acid sequence shown in SEQ ID NO. 2.		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

## PROTEIN TYROSINE KINASE AND LIGANDS THEREOF

5 The present invention relates to a putative signalling molecule. More particularly the invention relates to a novel tyrosine kinase and the use of this tyrosine kinase in the development of diagnostic and therapeutic approaches to cancer, for example breast cancer.

10 Protein tyrosine kinases are enzymes which show the property of catalysing the transfer of phosphate groups from donor molecules (ATP) to the hydroxyl groups of tyrosine residues in polypeptides. Known tyrosine kinases can be classified into two broad groups. Transmembrane tyrosine  
15 kinases traverse cellular membranes so that they possess extracellular and intracellular domains. Cytoplasmic tyrosine kinases are located only intracellularly. A general feature of transmembrane (also referred to as receptor) tyrosine kinases is that they possess extracellular ligand-binding  
20 domains, hydrophobic transmembrane sequences, and intracellular portions which include the tyrosine kinase domains (for a review see Ullrich & Schlessinger Cell, 61, 203-212 (1990)).

25 Tyrosine kinases can induce cell proliferation, cell transformation and regulate developmental events (see reviews by Hanks et al, Science, 241, 42-75 (1988) and Cantley et al, Cell 64, 281-302 (1991)). In general, where functional assays have been available (e.g. mitogenesis or transformation), it  
30 has been shown that the biological functions of tyrosine kinases are usually dependent on intact enzyme activity and that through autophosphorylation and the phosphorylation of other proteins, they alter the subcellular localisation and activities of various components of the intracellular  
35 signalling pathways.

Analysis of the oncogenes of many acutely transforming animal retroviruses has revealed that their products

SUBSTITUTE SHEET (RULE 26)

frequently manifest tyrosine kinase activity, as do the products of their cellular proto-oncogene counterparts. Other genes encoding tyrosine kinases have been found to be altered by DNA rearrangements in cancer cells, the result being a presumed acquisition of cellular transforming activity (for example c-abl, c-met). Some cellular proto-oncogenes encoding tyrosine kinases have been cloned independently by virtue of the fact that they encode growth factor receptors, for example the epidermal growth factor receptor. On the other hand, the genes for other growth factor receptors which have tyrosine kinase activities, such as those for platelet-derived growth factor and insulin-like growth factors, have been well characterised, but have never been found to be transduced by retroviruses.

The c-erbB-2/HER2/c-neu gene encodes a transmembrane receptor-like tyrosine kinase which is structurally very similar to the EGF/TGF alpha receptor. The c-erbB-2 gene has been found to be overexpressed in 20 to 30% of human breast tumours, often in association with gene amplification, and this phenotype is now generally accepted as predictive of poor disease free and overall survival (for reviews see Sunderland & McGuire in *Regulatory Mechanisms in Breast Cancer*, Lippman & Dickson (Eds.), Kluwer Academic Publishers, Boston, pages 3 to 22 (1991) and Gusterson et al., *J. Clin. Oncol.*, 10, 1049-1056 (1992)). Overexpression of the gene in fibroblasts induces transformation (di Fiore et al, *Science*, 237, 178-182 (1987) and Hudziak et al, *Proc. Natl. Acad. Sci. (USA)*, 84, 7159-7162 (1987)), and ligands which bind to the c-erbB-2 gene product (human and rodent) and activate its tyrosine kinase activity have been identified (Lupu et al, *Science*, 249, 1552-1555 (1990), Dobashi et al, *Proc. Natl. Acad. Sci. (USA)*, 88, 8582-8586 (1991), Wen et al, *Cell*, 69, 559-572 (1992) and Holmes et al, *Science* 256, 1205-1210 (1992)). Some of these ligands increase the rate of proliferation of cells expressing the c-erbB-2 protein. The strong implication of these and other findings is that overexpression of this tyrosine kinase in some breast tumour cells is one of the important steps in

their progression towards tumourigenicity, and therefore that c-erbB-2 can function as an oncogene.

5 A variety of other receptor tyrosine kinases, including those for the IGFs and the FGFs, are expressed in breast tumours (Stewart et al, J. Biol. Chem., 265 21172-21178 (1990) and Wellstein & Lippman in Molecular Foundations of Oncology, Broder (Ed.), Williams and Wilkins, Baltimore, pages 403-418 (1991)), and the EGF/TGF alpha receptor is overexpressed in  
10 some cases (for example Horak et al, Oncogene, 6, 2277-2284 (1991)). As with c-erbB-2, overexpression of EGFr has been correlated with poor prognosis. Since the ligands of some of these receptors can be detected in breast tumour samples, it may be that breast tumour development is also regulated by  
15 these molecules (for a review see van de Vijver & Nusse, Biochim, Biophys. Acta, 1072, 35-50 (1991)).

The features which are ideally required in a molecular target for tumour therapies include preferential expression  
20 in the tumours, a role in the development of the tumours (rather than merely being a marker), and a knowledge of the mechanism of its action as a starting point for the rational design of activity modulators. It appears that all of these three criteria are satisfied, at least to some extent, in the  
25 case of c-erbB-2 and breast cancer and a considerable research effort is currently being directed to this molecule.

It has recently been found that the alkaloid K252a, at appropriate concentrations, selectively inhibits the tyrosine  
30 kinase and biological activities of the trk-class of neurotrophin receptors, but has no effect on the activities of v-src, v-fms or the receptors for EGF and PDGF, and also does not affect general cell viability or proliferation (Tapley et al, Oncogene, 7, 371-381 (1992)). Members of  
35 another class of molecules, the tyrphostins, may be selective inhibitors of the EGFr/c-erbB-2 class of tyrosine kinases (Gazit et al, J. Med. Chem., 34, 1896-1907 (1991)). An antibody which inhibits the proliferation of human tumour

SUBSTITUTE SHEET (RULE 26)

cells by binding to the c-erbB-2 protein is currently under evaluation as a potential cancer therapy (Carter et al, Proc. Natl. Acad. Sci. (USA), 89, 4285-4289 (1992)).

5        If the transformation of a significant minority of breast tumour cells is due, at least in part, to the disregulated activity of the c-erbB-2 tyrosine kinase, then it seems likely that other breast tumours owe aspects of their transformed phenotype to the activity of other tyrosine kinases. The  
10       identification of further tumour associated tyrosine kinases would allow the development of novel diagnostic and therapeutic approaches to cancer and in particular breast tumours.

15       An object of the present invention is to provide a novel tyrosine kinase and its use in the diagnosis and/or therapy of cancer, particularly human breast tumours.

20       The present invention provides a DNA isolate encoding a novel protein tyrosine kinase or a fragment thereof, the protein tyrosine kinase being characterised by the amino acid sequence shown in

SEQ ID NO. 2

25       or by an amino acid sequence showing a significant degree of homology thereto, preferably at least 60% homology in the catalytic domain and at least 40% homology in the remainder of the molecule, more preferably at least 90% homology overall, for example at least 95% homology.

30       According to one embodiment, the DNA isolate encodes the full amino acid sequence of the novel protein tyrosine kinase or an amino acid sequence showing a significant degree of homology thereto. According to another embodiment the DNA  
35       isolate encodes a fragment of the amino acid sequence of the protein tyrosine kinase or an amino acid sequence showing a significant degree of homology thereto. DNA sequence encoding fragments of the protein tyrosine kinase preferably encodes those parts of the amino acid sequence which characterise the

enzyme, i.e. those parts which are most distinct from other protein tyrosine kinases.

5 The DNA sequence and deduced amino acid sequence of the novel protein tyrosine kinase are shown in SEQ ID NO. 1. Most preferably the DNA sequence encodes all or part of the catalytic domain of the enzyme. The catalytic domain is encoded by nucleotides 617-1324 of SEQ ID NO. 1.

10 The DNA isolate may have the base sequence defined in SEQ ID NO. 1 to encode the relevant amino acid sequence. Alternatively the DNA isolate may have any other DNA sequence encoding the relevant amino acid sequence.

15 According to another embodiment, the invention provides a DNA isolate encoding amino acids 1 to 77 of SEQ ID NO. 2 followed by the amino acids of SEQ ID NO. 12. This coding sequence of this isolate may have the base sequence of nucleotides 29 to 552 inclusive of SEQ ID NO. 1, but excluding  
20 nucleotides 259 to 380 inclusive, or any other base sequence encoding the relevant amino acid sequence.

The present invention also provides a DNA isolate in the form of a cloning vector or an expression vector preferably  
25 a plasmid vector, including DNA as defined above. In the case of an expression vector the DNA will be under control of an appropriate promoter and will include regulatory elements required for expression in a suitable host cell.

30 The present invention also includes a recombinant cell line transformed with an expression vector as defined above and capable of expressing a recombinant protein tyrosine kinase or a fragment thereof characterised by the amino acid sequence shown in

35 SEQ ID NO. 2 or an amino acid sequence showing a significant degree of homology thereto.

Suitable host cells include mammalian cells, insect cells, yeast cells and bacterial cells and the expression vector will include a promoter and other regulatory elements appropriate to the host cell in question. Preferred host  
5 cells include CHO cells, myeloma cells, immortalised human breast cells, rodent fibroblast cell lines, baculovirus cells, yeast and *E. coli* cells. General techniques for manipulating DNA coding sequences and expressing such sequences in various  
10 types of cells are well known to those skilled in the art and are described for example in Sambrook et al, Molecular Cloning: A Laboratory Manual, 2nd Edition, Cold Spring Harbor Laboratory.

The present invention also provides a protein tyrosine  
15 kinase or a fragment thereof characterised by the amino acid sequence shown in

SEQ ID NO. 2

or an amino acid sequence showing a significant degree of  
homology thereto.

20

The protein tyrosine kinase or a fragment thereof may be a recombinant polypeptide produced as the expression product of a coding sequence as defined above in a recombinant cell line. Alternatively the protein tyrosine kinase may be  
25 produced in a human tumour cell line. In either case the protein may be extracted and purified by standard techniques, for example antibody affinity chromatography. The glycosylation of the protein (if any) will depend on the cells in which it is produced and protein produced in human tumour  
30 cells in culture should have glycosylation equivalent to that produced *in vivo* in human tumours.

The DNA sequence set out in

SEQ ID NO. 1

35 represents DNA sequence and deduced amino acid sequence from a novel tumour protein tyrosine kinases.

The DNA sequences encoding the novel protein tyrosine



kinase has been isolated in the manner described in more detail in the example set out below but which involves briefly the following steps:

- i) isolation of mRNA from metastatic tumour tissue;
- 5 ii) preparation of cDNA from the isolated mRNA;
- iii) PCR amplification using degenerate oligonucleotide primers designed to amplify sequence associated with protein tyrosine kinases;
- iv) subcloning PCR products;
- 10 v) identification of protein tyrosine kinase products amplified at higher levels from tumour bearing tissue;
- vi) DNA sequencing and identification of novel protein tyrosine kinase sequence;
- vii) isolation of the putatively full length cDNA from a cDNA  
15 library.

Application of the above protocol identified the above novel protein tyrosine kinase. The enzyme, which is believed to be a cytoplasmic tyrosine kinase, has been found to be  
20 overexpressed in certain human breast tumours as compared to normal breast tissue but has not been detected at all in other types of human tissue.

The present invention also relates to the application of  
25 the novel tyrosine kinase for the development of therapeutic, prognostic and diagnostic approaches to cancer. The invention is particularly applicable to breast cancer, however, the same approaches may also be applicable to other cancers and research data suggests a connection between breast cancer and  
30 ovarian cancer and also between these two cancers and gastric cancer.

In terms of therapy, the involvement of the novel tyrosine kinase in tumours means that beneficial clinical  
35 effects in the treatment of tumours can be obtained by modulating the tyrosine kinase activity and/or the functionality of the signalling molecule. There are a number of ways in which such modulation could be achieved.

The activity of the protein tyrosine kinase could be significantly impaired or inhibited by small chemical molecules and screens to identify suitable small molecular weight inhibitors can be developed in more detail below. This approach is analogous to that proposed for other tyrosine kinases (see "Drugs of the Future", 17(2), 119-131, (1992)). A further embodiment of this approach would be to derive nmr spectra or crystal structures for the protein tyrosine kinases or domains thereof and use the structural information so obtained to synthesise chemical structures de novo which could be similarly screened for activity as inhibitors.

The approach described above could be extended to encompass small peptides which are either competitive for signalling action of the tyrosine kinase, or which demonstrate useful binding thereto, such that its function is inhibited. This would include, for example, peptides which are capable of blocking substrate binding to the signalling molecule, including but not limited to those which bear SH2 domains. An additional approach is to use small molecules or peptides to block or interfere with ligand activation or dimerisation of the tyrosine kinase. Both of these events are generally required to effect a biological signal so that blocking or inhibition thereof could be used to therapeutic effect.

Protein tyrosine kinase function could also be blocked by expressing or administering a peptide which is a truncated or altered version of the protein. These are typically referred to as dominant negative proteins and are believed to sequester the active protein in a non-functional complex. Protein tyrosine kinase function could also be modulated by blocking translation of the mRNA encoding the protein using antisense oligonucleotides.

A further regulatory role of the protein tyrosine kinase could be manipulated by controlling expression or activity of specific phosphatases which control the degree of protein phosphorylation. It is recognised that the phosphorylation

of specific residues, particularly tyrosine, threonine and serine residues, play an important role in transmitting a biological signal from the tyrosine kinase.

5        Use of the above therapeutic approaches depends on the identification and/or development of suitable agents for modulating, for example inhibiting, particular functions or activities of the protein tyrosine kinase. The present invention also relates to assays or screens by which such  
10       agents can be identified.

      The present invention also provides a method of screening a substance for potential utility as a therapeutic agent in the treatment of cancer, in particular breast cancer, which  
15       comprises providing a standard system in which a protein tyrosine kinase as defined above or an active fragment thereof is able to develop a measurable effect, allowing the protein tyrosine kinase to develop that effect in the presence and absence of the said substance and measuring that effect,  
20       ability to produce significant inhibition of the effect being taken as an indication of potential utility as a therapeutic agent.

      In accordance with this general principle, screens for  
25       potential therapeutic agents can take a number of forms. For example an initial screen to determine whether or not a substance merits further investigation as a potential inhibitor of the protein tyrosine kinase will usually be biochemical and should preferably be simple, rapid and capable  
30       of high through-put. Such screens will often make use of the protein expressed in a recombinant or derived from a cellular source in which the protein is over expressed. Particularly suitable recombinant expression systems include insect baculovirus, and the use of immortalised human breast cell  
35       lines or rodent fibroblast cell lines as host cells. The protein is then used directly in an *in vitro* assay with and without the potential inhibitor. This gives rapid data on the performance of the substance as an inhibitor often providing

direct evidence that the substance can inhibit the enzymic activity of the protein tyrosine kinase and thus merits further study. This type of isolated system can also provide data on specificity but will not generally provide any information on bioavailability. The enzyme activity of the protein tyrosine kinase *in vitro* can be measured, for example, by measuring autophosphorylation or phosphorylation of a model substrate, by following incorporation of radioactive phosphate or by using anti-phosphotyrosine antibodies.

Further information concerning potential inhibitors can be provided by cell based screens which make use of a phenotypic alteration, e.g. a change in morphology and/or tumorigenicity, conferred by expression of the protein tyrosine kinase in a recombinant cell line or over-expression in any other available cell line. Particularly suitable recombinant cell lines again use immortalised human breast cell lines or rodent fibroblast cell lines as host cells. The cells can be used in proliferation or tumorigenicity assays with and without the potential inhibitor looking for substances able to halt growth of and/or morphologically de-transform the cells. A control for substances which are generally cytotoxic can be provided by the same parental cell line engineered to express another oncogene which transforms the cells at a signalling point downstream of the protein tyrosine kinase. Assays of this sort can be very informative and they may provide data on mode of action of inhibitors.

The final stage of screening is the development of animal screens. These are time consuming and expensive and for these and ethical reasons their use is kept to a minimum, however they can give vital information regarding metabolism, clearance and performance of a substance in a true *in vivo* system which cannot be obtained in other ways. Animal screens ideally use the same recombinant or other cell lines as described above which express the protein tyrosine kinase. The cells are used as xenografts in animals, for example, nude mice, to give tumours whose growth is controlled by expression

of the protein tyrosine kinase. The mice are then given inhibitors to determine whether they are capable of causing the selective regression of the xenografts without blocking the growth of a control xenograft. Alternatively transgenic  
5 animals which develop tumours by virtue of the appropriate tissue specific expression of the target protein tyrosine kinase can be used in a similar manner.

The invention also extends to therapeutic agents  
10 identified by use of any or all of the screens referred to above. Preferably the therapeutic agent is a chemical molecule of relatively low molecular weight, for example, less than about 1000. Examples of suitable classes of molecule include staurosporine analogues, tyrphostins and flavenoids. The  
15 therapeutic agent can also be peptide or an antisense mRNA encoding the protein tyrosine kinase.

The invention also extends to an enzyme-substrate complex which comprises a protein tyrosine kinase characterised by the  
20 amino acid sequence shown in

SEQ ID NO. 2

or by an amino acid sequence showing a significant degree of homology thereto and a therapeutic agent capable of modulating the activity of the said protein tyrosine kinase.  
25

The novel protein tyrosine kinase can also be used in prognostic and diagnostic applications. Thus antibodies raised against the protein tyrosine kinase or nucleic acid encoding the novel tyrosine kinase can be used as the basis  
30 for screening tissue, in particular tumour tissue, for the presence of the novel protein tyrosine kinases. If, as in the case of c-erbB-2, the expression of the tyrosine kinase correlates with poor prognosis, then treatment of the patient could be modified accordingly.

35

Accordingly the present invention provides a method for detecting a tyrosine kinase having an amino acid sequence as defined in

## SEQ ID NO. 2

or a fragment thereof which comprises reacting a test sample with a specific antibody raised against an antigen from the said amino acid sequence and determining whether there is any  
5 antigen-antibody binding within the test sample. The test sample may be for example a tissue sample, such as a tumour sample and presence of abnormal amounts of the protein tyrosine kinase may indicate development of or susceptibility to develop a tumour.

10

The invention also provides a method for the detection in a sample of a nucleic acid sequence encoding a tyrosine kinase which comprises subjecting the sample or nucleic acid isolated therefrom to a method capable of detecting a nucleic  
15 acid sequence therein identical or substantially homologous to a defined nucleic acid sequence and identifying any sequence identical or substantially homologous to all or part of the DNA sequence defined in

## SEQ ID NO. 1,

20 DNA sequence from another part(s) of the tyrosine kinase characterised by that sequence or mRNA which would be the transcription product of such DNA. Again presence of abnormal amounts of nucleic acid encoding the protein tyrosine kinase may indicate development of or susceptibility to develop a  
25 tumour.

Suitable methods for the detection of specific DNA sequences include Southern blotting and/or PCR using appropriate primers. Suitable methods for detecting mRNA  
30 include Northern blotting, RNase protection studies and direct nucleic acid *in situ* hybridisation.

Experimental work on which the invention is based is described in more detail in the following Example in which  
35 reference is made to the following figures in which:

Figure 1 shows expression of brk transcripts in human tissues and breast tumour cell lines;

13

Figure 2 shows a schematic representation of the human brk cDNA;

Figure 3 shows baculovirus and bacterial protein expression of brk;

- 5 Figure 4 shows detection of brk mRNA expression in normal breast and grade III breast carcinoma tissue by RNase protection;

Figure 5 shows detection of brk mRNA expression in normal breast and grade III breast carcinoma tissue by PCR.

10

SUBSTITUTE SHEET (RULE 26)

EXAMPLEMaterials and Methods5    **Tissue samples**

Human lymph node and tumour samples were obtained from surgery and rapidly frozen in liquid nitrogen. Tumour grading was in accordance with accepted national criteria (Pathology Reporting in Breast Cancer Screening, Draft Guidance. Prepared by a Department of Health and Royal College of Pathologists Working Group (1989)) Breast Screening Publication ISBN 187997101. Normal breast tissue was obtained from reduction mammoplasty and rapidly frozen in liquid nitrogen. Frozen tissue samples were stored either in liquid nitrogen or at -70°C.

**Cell lines**

20        Breast tumour cell lines were obtained from the American Type Culture Collection. Cell lines were maintained under conditions recommended by the supplier.

**RNA purification and Northern hybridisation**

25        Total RNA was purified from tissue samples using the method of Chomczynski and Sacchi (Anal. Biochem., 162, 156-159 (1987)). Cytoplasmic RNA was extracted from cell lines by the method of Wilkinson (Nuc. Acids. Res., 16, 10933 (1988)).

30    Poly A<sup>+</sup> mRNA was purified using oligo dT spun columns (Pharmacia) according to manufacturer's instructions. RNA samples were electrophoresed through formaldehyde-agarose (1%) gels and transferred to Hybond N membranes (Amersham). The human tissue poly A<sup>+</sup> Northern blot was purchased from

35    Clontech. <sup>32</sup>P radiolabelled probes were generated as previously described (Feinberg and Vogelstein, Anal. Biochem., 132, 6-13 (1983)) and hybridised to membranes according to manufacturer's instructions. Membranes were washed at 65°C



15

at a final stringency of 0.1 x SSC, 0.1% SDS.

**Reverse transcriptase-polymerase chain reaction (RT-PCR), and  
PTK differential screen**

5

For the production of single stranded cDNA, 5 µg of RNA were digested with 7.5 units (u) RNase-free DNase I in 10 µl for 10 mins. at 37°C in the presence of 40u human placental RNase inhibitor. After extraction with phenol/chloroform and chloroform, the RNA was precipitated, resuspended in 10 µl water, denatured at 65°C for 5 mins., and stored on ice. Reverse transcription reactions (10 µl) contained 2.5 µg RNA, 40 ng pdN6 (Pharmacia), 1mM dNTPs, 40u RNase inhibitor, 30u AMV reverse transcriptase XL (Life Sciences Inc.) and reverse transcriptase buffer produced by the enzyme supplier. Control reactions contained no reverse transcriptase. Reactions were carried out at 41°C for 1 hour.

Degenerate deoxyoligonucleotides were synthesised to amplify PTK encoding cDNA sequences.

SEQ ID NO. 3 (sense)

SEQ ID NO. 4 (antisense)

These primers are based on the following amino acid sequences

SEQ ID NO. 5

25 SEQ ID NO. 6

respectively.

Amplifications were carried out in 50 µl volumes containing 2 µl of single stranded cDNA, 60 mM KCl, 15 mM Tris-HCl (pH 8.8), 1.1 mM MgCl<sub>2</sub>, 200 µM of each dNTP, 20 pmol of each primer and 2u of Taq polymerase (Cetus). The PCR temperature cycle parameters were: 94°C for 1 minute (min), 50°C for 2 mins, 72°C for 1 min, 30 cycles. For each RNA sample, control amplifications were performed on samples which had not subjected to reverse transcription (RT-controls).

The amplified PTK PCR products were electrophoresed through 3% agarose gels and purified using "Mermaid" reagents

SUBSTITUTE SHEET (RULE 26)

(Bio 101 inc). PCR products were subcloned into the vector PCR II (Invitrogen) according to supplier's instructions. Bacterial cell lysates from the involved node PTK fragment library were spotted on to duplicate Hybond N membranes. PTK  
5 PCR products amplified at higher levels from tumour bearing lymph nodes with respect to non-involved nodes were identified by differential DNA hybridisation, using the protocol of Crompton et al, Nacl. Acids Res., 20, 4107-4108 (1992).

10 RT-PCR expression studies were carried out using specific primers based on the brk  $\lambda$ t2 cDNA sequence (sense: 485-502, antisense: 1008-989) and the human glyceraldehyde 3-phosphate dehydrogenase (GAPDH) cDNA sequence (Arcari et al., Nacl. Acids Res., 12, 9179-9189 (1984)) (sense:71-96, antisense  
15 1030-1053). Reverse transcription and PCR were carried out as described above, with all four primers in the same reaction. RT-controls were performed for each sample. After electrophoresis through 3% agarose gels, amplified GAPDH products were detected by ethidium bromide staining.  
20 Amplified brk products were detected by Southern transfer to Hybond-N membranes (Amersham) and probing with radiolabelled  $\lambda$ t2 cDNA sequences (Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press (1989)).

25

#### RACE PCR

CDNA sequences 3' of the region amplified by the PTK PCR were isolated by a modification of the RACE protocols  
30 described by Frohman et al. (Proc. Natl. Acad. Sci. (USA), 85, 8998-9002 (1988)). T-47D RNA was reverse transcribed using a 'tagged' oligo dT primer as shown in

SEQ ID NO. 7.

Two specific sense oligodeoxynucleotides were used for the brk  
35 kinase sequence in consecutive PCR reactions. These were as shown in

SEQ ID NO. 8

SEQ ID NO. 9.

The reaction buffer was the same as described above except the  $MgCl_2$  concentration was 1.5 mM. The PCR cycle parameters consisted of a single denaturation step at 92°C for 2 mins, followed by 30 cycles at 92°C for 1.25 mins, 56°C for 2 mins and 72°C for 2 mins.

5' RACE PCR was carried out using the Life Technologies 5' RACE system according to manufacturer's instructions. The two nested brk specific antisense oligodeoxynucleotides were as shown in

SEQ ID NO. 10

SEQ ID NO. 11.

The PCR cycle parameters consisted of a single denaturation step at 92°C for 2 mins followed by 30 cycles at 92°C for 1 min, 55°C for 1.5 mins and 72°C for 1.5 mins.

#### Constructions and screening of cDNA libraries

Cell line cDNA libraries were prepared in the vector lambda ZAPII (Stratagene) using reverse transcriptase (Superscript II, Life Technologies) and a unizAP-cDNA synthesis kit (Stratagene). The complexities of the T-47D and MCF-7 libraries were  $2.1 \times 10^6$  and  $5 \times 10^6$ . Libraries were screened using Hybond N nylon membranes (Amersham). <sup>32</sup>p radiolabelled DNA probes were hybridised to the membranes according to manufacturer's instructions. Filters were washed in 2 x SSC, 0.1% SDS at 65°C and finally in 0.1 x SSC, 0.1% SDS at 65°C.

#### DNA sequencing

cDNA inserts were sequenced by the dideoxy method (Sanger et al., 1977) using the Sequenase version 2 sequencing kit (United States Biochemicals) after generating exonuclease III nested deletions in the phagemids SK<sup>+</sup> and KS<sup>+</sup> (Stratagene) and rescuing single stranded templates as described by the supplier.

### Baculoviral expression of brk

The brk cDNA  $\lambda$ t2 was subcloned into the baculovirus transfer vector pVL1392 (Invitrogen). This was co-transfected with BaculoGold (Pharmingen) viral DNA onto Spodoptera frugiperda (Sf21) insect cells and three days later the medium was used in an agarose plaque assay. Recombinant plaques containing brk were isolated, characterised, and used to generate high titre viral stocks ( $>3 \times 10^8$  pfu/ml). Monolayer cultures of Sf-21 cells were subsequently infected at a multiplicity of infection of 5 pfu/cell and analysed after 48 hrs. Insect cell lysates were prepared as previously described (Page et al., J. Biol. Chem., 264, 19147-19154 (1989)).

### Bacterial expression of brk and immunodetection

A SmaI-HindIII fragment from clone  $\lambda$ t2, encoding amino acids 3-451, was cloned into the bacterial expression plasmid pGEX-4T-3 (Pharmacia). Bacterial cultures expressing this plasmid were grown at 37°C in L-broth containing 100  $\mu$ g/ml ampicillin and induced with 0.1 mM isopropyl- $\beta$ -D-thiogalactopyranoside for 2 hours. Induced bacteria were lysed by boiling in SDS-PAGE sample buffer.

Cell lysates were separated on 10% SDS-polyacrylamide gels and transferred to nitrocellulose membranes. Membranes were blocked by incubation in phosphate-buffered saline (PBS) containing 5% (w/v) non-fat dried milk-protein, washed with PBS/0.05% Triton X100, and incubated with the anti-phosphotyrosine monoclonal antibody 4G10 (UBI) (2 ng/ml) in L15 medium (Gibco BRL) containing 10% (v/v) FCS. The antibody was detected using the ECL detection reagents (Amersham) as according to manufacturer's instructions.

Construction of expression vectors for expression of brk in mammalian cells

5 The brk cDNA  $\lambda$ t2 was isolated from the cDNA library as a subclone between the EcoRI and XhoI sites of pBluescript SK (-) (Stratagene). The cDNA was excised by digestion with KpnI and BamHI, its ends were blunted, and it was subcloned into the SmaI site of pBluescript KS (+) (Stratagene). Ligation in one orientation produced a construct whereby digestion with  
10 HindIII produced a fragment carrying the full brk coding sequence. This HindIII fragment was ligated into either pRCCMV (Invitrogen) or pREP8 (Invitrogen), and in each case clones with the 'sense' orientation were selected.

15 Stable transfection of human mammary epithelial cells

The mammary epithelial cell line HB4a has shown itself to be resistant to transfection and a modified calcium phosphate transfection technique was used specifically for  
20 these cells to ensure successful stable expression of the vector. This was achieved using an episomal vector to provide stable expression of the brk cDNA and uniquely demonstrate a transformed phenotype. Two parameters proved critical to achieving increased transfection efficiency:

25 1. Use of OPTIMEM, an  $\alpha$ -MEM based medium (Gibco), for seeding and transfection. A special feature of this medium is reduced phenol red (a pH indicator dye which is present in all media) and it is believed that the dye may interfere with transfection of HB4a cells.

30 2. A specific coarseness of calcium phosphate-DNA complex achieved by mixing DNA and calcium chloride with immediate vortexing for a duration of six seconds. The cells were then exposed to the precipitate overnight, shocked with 15% glycerol for exactly 2.5 min and then returned to standard  
35 growth medium.

### RNase protection

A Pst I fragment of the  $\lambda$ t2 brk cDNA (bases 302-607) was subcloned into the vector pBluescript (Stratagene). The pTRI-  
5 GAPDH vector, containing a 316 bp human glyceraldehyde 3-phosphate dehydrogenase cDNA fragment, was obtained from Ambion Inc. Radiolabelled antisense RNA probes were transcribed from linearised vectors using standard procedures (Sambrook et al., Molecular Cloning: A Laboratory Manual,  
10 Cold Spring Harbor Laboratory Press (1989)). Probe purification and RNase protection were carried out using protocols and reagents in the 'Ambion RPA II' kit. Protected fragments were resolved on DNA sequencing polyacrylamide/urea gels, and detected in dried gels by autoradiography.

15

### RESULTS

#### Isolation of PTK cDNAs differentially expressed in a breast tumour metastasis

20

RNA was isolated from involved and non-involved axillary nodes removed from a patient with metastatic breast carcinoma. The RNA was subjected to reverse transcription-polymerase chain reaction (RT-PCR) using degenerate oligonucleotides  
25 corresponding to the highly conserved amino acid sequences present in subdomains VI (SEQ ID NO. 5) and IX (SEQ ID NO. 6) of the PTK catalytic domain (Hanks et al., Science, 241, 42-52 (1988)). The pool of PTK cDNA fragments amplified from involved node RNA was subcloned to create a plasmid library  
30 of PTK cDNAs. Radiolabelled probes prepared from the two pools of PTK fragments amplified from the involved and non-involved node RNAs were used to screen duplicate platings of the PCR fragment library. Of 50 duplicate clones screened, 4 were found to hybridise preferentially to the probe prepared  
35 from the PTK cDNA fragment pool amplified from the involved node RNA. Sequencing of these clones identified three different cDNA fragments whose deduced amino acid sequences inferred that they encoded PTKs. One of the cDNA fragments

(designated PTK23) encoded an amino acid sequence showing homology to members of the src PTK family blk, c-src and c-yes. Further characterisation of this novel putative PTK named brk (breast tumour kinase) is described herein.

5

#### Isolation of brk cDNAs

Initially, a 600 base pair (bp) brk cDNA fragment was generated by 3' RACE (rapid amplification of cDNA ends) PCR, using two specific nested oligonucleotides corresponding to the brk sequence between the two degenerate primers. This 3' RACE fragment was then used in expression studies to identify human breast tumour cell lines and normal tissues that expressed the brk transcript. Northern analysis using the 600 bp fragment on total RNA isolated from a range of human breast tumour cell lines (MCF-7, T-47D, MDA MB 468, -415, -157 and BT474, data not shown) did not detect any brk transcripts. Using the same fragment as a probe on Northern blots containing poly A<sup>+</sup> selected RNA, two transcripts in the size range of 2.3 kb and 2.6-2.8 kb were detected in the two human breast tumour cell lines tested (T-47D and MCF-7).

Figure 1 shows expression of brk transcripts in human tissues and breast tumour cell lines. Northern blots, containing 2 µg of poly A<sup>+</sup> RNA isolated from human tissues and two human breast tumour cell lines, were hybridised with a human brk probe (upper panel) and GAPDH probe (lower panel). The blots subjected to autoradiography for 10 days (brk) and 36 hours (GAPDH). The mobilities of the molecular weight markers are indicated in kb.

Subsequently, the 600 bp fragment was used to screen an unamplified cDNA library prepared from the T-47D cell line. From a screen of  $1 \times 10^6$  phage clones, two brk cDNAs were isolated, both of which contained inserts of approximately 2.5 kb.

Figure 2 shows a schematic representation of the human

brk cDNA showing selected restriction sites, the relative positions of the end of the cDNA clones ( $\lambda$ t1,2 and  $\lambda$ m3-5) and the positions of the 5' and 3' RACE-PCR generated fragments. The shaded region represents the open reading frame present in clone  $\lambda$ t2. P1 represents the 0.3 kb PstI-PstI fragment used in the RNase protection assay.

Sequencing of the slightly longer clone  $\lambda$ t2 identified an open reading frame (ORF) of 1353 bp starting with a consensus sequence for efficient translational initiation (Kozak, J. Cell. Biol., 108, 229-241, (1989)) and terminating with a TGA codon at position 1392. The ORF was preceded by a 5' sequence of 28 bp that did not contain any in-frame stop codons.

The possibility arose that the cDNA clone  $\lambda$ t2 did not contain the complete brk coding region. In order to identify potentially longer cDNAs, a radiolabelled restriction fragment corresponding to the 5' end of clone  $\lambda$ t2 was used as a probe to screen  $1 \times 10^6$  phage clones from a MCF-7 cell line unamplified cDNA library. From this screen three cDNAs ( $\lambda$ m3-5) were isolated (Figure 2), one of which ( $\lambda$ m4) had an additional 17 bp of 5' sequence. No initiation or stop codons in frame with the ORF were identified in this upstream sequence. In addition to these studies, 5' RACE PCR was carried out using nested oligonucleotide primers, corresponding to the 5' end of clone  $\lambda$ t2, on poly A<sup>+</sup> RNA isolated from T-47D and MCF-7 cell lines. Subcloning and sequencing of the amplified products indicated that no more upstream sequence had been isolated (Figure 2). Expression of clone  $\lambda$ t2, by recombinant baculovirus infection of Sf21 cells, resulted in the synthesis of a protein whose mobility on SDS-polyacrylamide gels was consistent with the predicted molecular weight of the protein encoded by the ORF present.

Figure 3 shows baculoviral and bacterial protein expression of brk. Lysates from insect cells infected with wild type baculovirus (AcPNV) and a recombinant baculovirus



containing the brk cDNA  $\lambda$ t2 were subjected to SDS-PAGE (panel A). Proteins were visualised by staining with Coomassie blue G-250. Arrows indicate the positions of the wild type polyhedrin protein and the recombinant brk protein. The same insect cell lysates (panel B) and lysates from bacteria containing plasmids expressing GST or GST-brk fusion proteins (panel C) were subjected to SDS-PAGE and transferred to nylon membranes. Phosphotyrosine containing proteins were detected using the antibody 4G10. Duplicate membranes were incubated with only the detection reagents to confirm that detection was dependant on the 4G10 antibody. The arrows indicate the positions of polyhedrin (panel B) and GST proteins (panel C), as determined by Ponceau S staining of the membranes. The mobilities of the molecular weight markers are indicated in kDa.

In addition, *in vitro* translation in rabbit reticulocyte lysates using transcripts generated from clone  $\lambda$ t2 resulted in the synthesis of a single protein with the same mobility (data not shown). These experiments indicated that the predicted start codon in the brk cDNAs could act as an initiation site for translation and, therefore, that clone  $\lambda$ t2 most likely contained the complete brk coding region.

## brk encodes a novel putative non-receptor kinase

The 1353 bp open reading frame present in cDNA  $\lambda$ t2 encodes a protein of 451 amino acids with a predicted molecular weight of 52 kDa. The complete nucleotide sequence and amino acid translation of  $\lambda$ t2 are shown in SEQ ID NO. 1. The amino acid sequence between positions 197-432 of brk has several motifs found in other PTK catalytic domains (Hanks et al., Science, 241, 42-52 (1988)). The putative ATP-binding motif, GSGYFG, is at position 198-203, with a downstream lysine residue at 219. The putative autophosphorylation site in the catalytic domain is conserved at position 342. In brk the strong indicator sequence of PTK specificity in subdomain VI corresponds to DLAARN (position 312-317). This motif is

shared with abl, fps, csk and bcr(atk), but not with the src subfamily genes (DLRAAN). Amino acid sequences similar to SH2 (78-169) and SN3 (15-68) domains are also present. The predicted amino terminal methionine of brk is followed by valine and residue 7 is alanine. brk therefore lacks the residues thought to be necessary for post-translational N-terminal myristoylation (Kaplan et al., Mol. Cell. Biol., 8, 2435-2441 (1988)). The short carboxyl terminus following the tyrosine kinase domain of brk contains a potential regulatory tyrosine (position 447) equivalent to tyrosine 527 of c-src, but it is found in a very different sequence context. The predicted amino-terminal region of brk, beyond the putative SH3 domain, is very short and is equivalent in length to that of csk (Nada et al, Nature, 351, 69-72 (1991); Partanen et al., Oncogene, 6, 2013-2018 (1991); Brauninger et al., Gene, 110, 205-211 (1992)).

#### Expression of brk in human tissues and breast tumour cell lines

As mentioned previously, brk transcripts were not detected when Northern blots containing total RNA isolated from a variety of breast tumour cell lines were screened with the 600 bp brk 3' RACE fragment. Two transcripts in the size range of 2.3 kb and 2.6-2.8 kb were detected on Northern blots containing poly A<sup>+</sup> mRNA isolated from the breast tumour cell lines MCF-7 and T-47D. No brk transcripts were detected on Northern blots containing poly A<sup>+</sup> mRNA isolated from normal human heart, placenta, brain, lung, liver, skeletal muscle, kidney and pancreas (Figure 1).

Using RNase protection and RT-PCR assays, we examined the expression of brk in normal breast samples (reduction mammoplasty) and in grade III breast carcinomas. In the RNase protection assay, a 305 bp protected fragment of similar intensity to that detected in T-47D RNA was detected in one out of five grade III breast carcinoma RNAs but not in two normal breast samples.

Figure 4 shows detection of brk mRNA expression in normal breast and grade III breast carcinoma tissue by RNase protection. Total RNA samples were isolated from yeast, T-47D cell line, two different normal human breast samples and five different grade III breast carcinomas. RNAs were hybridised with a brk radiolabelled antisense RNA derived from the P1 cDNA fragment (Figure 2) or with an antisense GAPDH RNA probe. After RNase digestion and separation on a sequencing gel, products were detected by autoradiography for 4 days at -70°C.

Figure 5 shows detection of brk mRNA expression in normal breast and grade III breast carcinoma tissue by PCR. Specific primers were used to simultaneously amplify brk and GAPDH sequences present in cDNA synthesised from normal breast (N) or grade III breast carcinoma (T) RNAs. Products were resolved on agarose gels, and the amplified GAPDH products were visualised with ethidium bromide (lower portion of each panel). After transfer to nylon membranes, brk products were detected by probing with a specific radiolabelled cDNA and autoradiography (upper portion of each panel). Lanes immediately to the left of the visualised GAPDH sequences contain the products of amplification from the corresponding non-reverse-transcribed RNAs. The upper panel corresponds to the same RNA samples as shown in Figure 4, excepting the left hand 'normal' sample in Figure 4 which is omitted here. The two lower panels represent an additional 4 normal and 7 tumour samples.

By RT-PCR, a 600 bp reverse transcriptase-dependent brk product was specifically amplified from the tumour RNA which gave a positive signal in the RNase protection assay, when the same tumour and normal breast RNAs were analysed (Figure 5, upper panel). Analysis of further normal breast and grade III breast carcinoma RNAs by RT-PCR detected another three out of seven tumours that expressed brk transcripts, whereas no brk products were amplified from four normal breast RNA samples (Figure 5, lower panels).

### Baculoviral and bacterial expression of brk cDNA

A recombinant baculovirus was generated containing the entire  $\lambda$ t2 cDNA under the control of the polyhedrin promoter. 5  
Infected insect cells expressed a novel protein with an SDS-PAGE mobility of approximately 50 kDa (Figure 3, A), consistent with the predicted molecular weight of 52 kDa. A bacterial vector was constructed to express the entire brk open reading frame, except for the first two N-terminal amino 10  
acids, as a glutathione S-transferase (GST) fusion protein. Cell lysates from bacterial cultures containing the GST-brk vector expressed an approximately 70 kDa fusion protein.

Western blots of both insect and bacterial cell lysates 15  
using an anti-phosphotyrosine antibody demonstrated that the baculovirally expressed 50 kDa protein and the bacterially expressed 70 kDa fusion protein were the major immunoreactive polypeptides present (Figure 3). No immunoreactive polypeptides were present in insect cell lysates infected with 20  
a wild type baculovirus and cell lysates from bacteria expressing GST alone (Figure 3).

### Expression and function of brk cDNAs in mammalian cells

25  
NIH3T3 (clone b2B) cells were transfected either with pRCCMV or its derivative carrying the brk cDNA. The cells were selected with 1 mg/ml Geneticin, and clones of drug resistant cells were pooled. These populations were seeded at either  $1 \times 10^4$  or  $1 \times 10^5$  per well of a six well dish in 0.4% 30  
agar, and left for six weeks. At the end of this period, 20-30  $\mu$ m colonies were observed with the cells containing the transfected brk cDNA, whereas none were observed with the cells transfected with the empty vector. In multiple seedings of  $1 \times 10^4$  'brk' cells, the numbers of colonies obtained were 35  
44, 54 and 0; in the  $1 \times 10^5$  seedings there were 81, 41, 50 and 42 colonies. Of 55 of these colonies, 32 were successfully isolated and grown on, whereas none of the cell aggregates picked from the 'vector only' cells in agar were capable of

further proliferation. Seven of the expanded 'brk' clones were chosen at random and placed back in soft agar; five of these produced 50  $\mu$ m colonies with 100% efficiency within seven days when seeded at  $1 \times 10^4$  (lower efficiencies were observed at higher seeding densities). These data indicate that brk expression in N1H3T3 cells results in anchorage independent proliferation, a feature of neoplastically transformed cells. The clones of brk transformed N1H3T3 cells will be useful tools for the study of brk function, and for screening potential brk inhibitors.

HB4a, a conditionally immortalised human breast luminal cell line (developed by Dr. M. O'Hare), was transfected with pREP8 or its derivative containing the brk cDNA, and transfectants were selected in 1mM histidinol. For each plasmid, 50-60 drug resistant colonies were pooled and maintained under selection. The cells were seeded at  $1 \times 10^5$  in duplicate in 0.4% agar and left for six weeks. The brk transfected cells produced an average of 373 50  $\mu$ m clones, whereas the empty vector cells produced an average of 108. Thus, brk expression in these cells enhances their capacity to proliferate in an anchorage independent manner.

#### An alternative brk cDNA

Sequencing of brk cDNA clone  $\lambda$ m5 revealed a deletion corresponding to nucleotides 259 to 380 inclusive of SEQ ID NO. 1. The predicted protein product is identical to that in SEQ ID NO. 2 from the N-terminus up to and including residue 77. Following the deletion, the nucleotide sequence continues to correspond to that shown in SEQ ID NO. 1 (nucleotides 381 onwards) but with the reading frame changed so that the amino acid sequence diverges. The ensuing predicted amino acid sequence is shown in SEQ ID NO. 12 which is followed by a stop codon.

This protein would, therefore, comprise an SH3 domain identical in sequence to that of brk, but lack the SH2 and catalytic domains. Since SH3 domains may mediate protein-

protein interactions, this protein may regulate brk function by competing for binding to specific protein partners.

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

## (i) APPLICANT:

(A) NAME: The Institute of Cancer Research  
 (B) STREET: 17A, Onslow Gardens  
 (C) CITY: London  
 (E) COUNTRY: Great Britain  
 (F) POSTAL CODE (ZIP): SW7 3AL

(A) NAME: The Wellcome Foundation Limited  
 (B) STREET: Unicorn House, 160 Euston Road  
 (C) CITY: London  
 (E) COUNTRY: Great Britain  
 (F) POSTAL CODE (ZIP): NW1 2BP

(ii) ~~TITLE OF INVENTION: Signalling Molecule~~

(iii) NUMBER OF SEQUENCES: 12

## (iv) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk  
 (B) COMPUTER: IBM PC compatible  
 (C) OPERATING SYSTEM: PC-DOS/MS-DOS  
 (D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

## (2) INFORMATION FOR SEQ ID NO: 1:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2507 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

## (ix) FEATURE:

(A) NAME/KEY: CDS  
 (B) LOCATION: 29..1381

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

45	CCTGGTCCTG CCGCTGCGCC CGCCCGCC ATG GTG TCC CGG GAC CAG GCT CAC	52
	Met Val Ser Arg Asp Gln Ala His	
	1 5	
50	CTG GGC CCC AAG TAT GTG GGC CTC TGG GAC TTC AAG TCC CGG ACG GAC	100
	Leu Gly Pro Lys Tyr Val Gly Leu Trp Asp Phe Lys Ser Arg Thr Asp	
	10 15 20	
55	GAG GAG CTG AGC TTC CGC GCG GGG GAC GTC TTC CAC GTG GCC AGG AAG	148
	Glu Glu Leu Ser Phe Arg Ala Gly Asp Val Phe His Val Ala Arg Lys	
	25 30 35 40	
60	GAG GAG CAG TGG TGG TGG GCC ACG CTG CTG GAC GAG GCG GGT GGG GCC	196
	Glu Glu Gln Trp Trp Trp Ala Thr Leu Leu Asp Glu Ala Gly Gly Ala	
	45 50 55	
65	GTG GCC CAG GGC TAT GTG CCC CAC AAC TAC CTG GCC GAG AGG GAG ACG	244
	Val Ala Gln Gly Tyr Val Pro His Asn Tyr Leu Ala Glu Arg Glu Thr	
	60 65 70	
	GTG GAG TCG GAA CCG TGG TTC TTT GGC TGC ATC TCC CGC TCG GAA GCT	292
	Val Glu Ser Glu Pro Trp Phe Phe Gly Cys Ile Ser Arg Ser Glu Ala	
	75 80 85	

SUBSTITUTE SHEET (RULE 26)

30

	GTG CGT CGG CTG CAG GCC GAG GGC AAC GCC ACG GGC GCC TTC CTG ATC Val Arg Arg Leu Gln Ala Glu Gly Asn Ala Thr Gly Ala Phe Leu Ile 90 95 100	340
5	AGG GTC AGC GAG AAG CCG AGT GCC GAC TAC GTC CTG TCG GTG CGG GAC Arg Val Ser Glu Lys Pro Ser Ala Asp Tyr Val Leu Ser Val Arg Asp 105 110 115 120	388
10	ACG CAG GCT GTG CGG CAC TAC AAG ATC TGG CGG CGT GCC GGG GGC CGG Thr Gln Ala Val Arg His Tyr Lys Ile Trp Arg Arg Ala Gly Gly Arg 125 130 135	436
15	CTG CAC CTG AAC GAG GCG GTG TCC TTC CTC AGC CTG CCC GAG CTT GTG Leu His Leu Asn Glu Ala Val Ser Phe Leu Ser Leu Pro Glu Leu Val 140 145 150	484
	AAC TAC CAC AGG GCC CAG AGC CTG TCC CAC GGC CTG CGG CTG GCC GCG Asn Tyr His Arg Ala Gln Ser Leu Ser His Gly Leu Arg Leu Ala Ala 155 160 165	532
20	CCC TGC CGG AAG CAC GAG CCT GAG CCC CTG CCC CAT TGG GAT GAC TGG Pro Cys Arg Lys His Glu Pro Glu Pro Leu Pro His Trp Asp Asp Trp 170 175 180	580
25	GAG AGG CCG AGG GAG GAG TTC ACG CTC TGC AGG AAG CTG GGG TCC GGC Glu Arg Pro Arg Glu Glu Phe Thr Leu Cys Arg Lys Leu Gly Ser Gly 185 190 195 200	628
30	TAC TTT GGG GAG GTC TTC GAG GGG CTC TGG AAA GAC CGG GTC CAG GTG Tyr Phe Gly Glu Val Phe Glu Gly Leu Trp Lys Asp Arg Val Gln Val 205 210 215	676
35	GCC ATT AAG GTG ATT TCT CGA GAC AAC CTC CTG CAC CAG CAG ATG CTG Ala Ile Lys Val Ile Ser Arg Asp Asn Leu Leu His Gln Met Leu 220 225 230	724
	CAG TCG GAG ATC CAG GCC ATG AAG AAG CTG CGG CAC AAA CAC ATC CTG Gln Ser Glu Ile Gln Ala Met Lys Lys Leu Arg His Lys His Ile Leu 235 240 245	772
40	GCG CTG TAC GCC GTG GTG TCC GTG GGG GAC CCC GTG TAC ATC ATC ACG Ala Leu Tyr Ala Val Val Ser Val Gly Asp Pro Val Tyr Ile Ile Thr 250 255 260	820
45	GAG CTC ATG GCC AAG GGC AGC CTG CTG GAG CTG CTC CGC GAC TCT GAT Glu Leu Met Ala Lys Gly Ser Leu Leu Glu Leu Arg Asp Ser Asp 265 270 275 280	868
50	GAG AAA GTC CTG CCC GTT TCG GAG CTG CTG GAC ATC GCC TGG CAG GTG Glu Lys Val Leu Pro Val Ser Glu Leu Leu Asp Ile Ala Trp Gln Val 285 290 295	916
55	GCT GAG GGC ATG TGT TAC CTG GAG TCG CAG AAT TAC ATC CAC CGG GAC Ala Glu Gly Met Cys Tyr Leu Glu Ser Gln Asn Tyr Ile His Arg Asp 300 305 310	964
	CTG GCC GCC AGG AAC ATC CTC GTC GGG GAA AAC ACC CTC TGC AAA GTT Leu Ala Ala Arg Asn Ile Leu Val Gly Glu Asn Thr Leu Cys Lys Val 315 320 325	1012
60	GGG GAC TTC GGG TTA GCC AGG CTT ATC AAG GAG GAC GTC TAC CTC TCC Gly Asp Phe Gly Leu Ala Arg Leu Ile Lys Glu Asp Val Tyr Leu Ser 330 335 340	1060
65	CAT GAC CAC AAT ATC CCC TAC AAG TGG ACG GCC CCT GAA GCG CTC TCC His Asp His Asn Ile Pro Tyr Lys Trp Thr Ala Pro Glu Ala Leu Ser 345 350 355 360	1108

SUBSTITUTE SHEET (RULE 26)



31

	CGA GGC CAT TAC TCC ACC AAA TCC GAC GTC TGG TCC TTT GGG ATT CTC	1156
	Arg Gly His Tyr Ser Thr Lys Ser Asp Val Trp Ser Phe Gly Ile Leu	
	365 370 375	
5	CTG CAT GAG ATG TTC AGC AGG GGT CAG GTG CCC TAC CCA GGC ATG TCC	1204
	Leu His Glu Met Phe Ser Arg Gly Gln Val Pro Tyr Pro Gly Met Ser	
	380 385 390	
10	AAC CAT GAG GCC TTC CTG AGG GTG GAC GCC GGC TAC CGC ATG CCC TGC	1252
	Asn His Glu Ala Phe Leu Arg Val Asp Ala Gly Tyr Arg Met Pro Cys	
	395 400 405	
15	CCT CTG GAG TGC CCG CCC AGC GTG CAC AAG CTG ATG CTG ACA TGC TGG	1300
	Pro Leu Glu Cys Pro Pro Ser Val His Lys Leu Met Leu Thr Cys Trp	
	410 415 420	
20	TGC AGG GAC CCC GAG CAG AGA CCC TGC TTC AAG GCC CTG CGG GAG AGG	1348
	Cys Arg Asp Pro Glu Gln Arg Pro Cys Phe Lys Ala Leu Arg Glu Arg	
	425 430 435 440	
	CTC TCC AGC TTC ACC AGC TAC GAG AAC CCG ACC TGAGCTGCTG TGGAGCGGGC	1401
	Leu Ser Ser Phe Thr Ser Tyr Glu Asn Pro Thr	
	445 450	
25	ATGGCCGGGC CCTGCTGAGG AGGGGCCTGG GCAGAGGGCC TGGACCTGGG ATCAAGGCCC	1461
	ACGCGCTTCC CTGGGGTTTA CTGAGGTGAT GGGTGCAGGA AAGGTTTACA AATGTGGAGT	1521
30	GTCTGCGTCC AATACACGCG TGTGCTCCTC TCCTTACTCC ATCGTGTGTG CCTTGGGTCT	1581
	CAGCTGCTGA CACGCAGCCT GCTCTGGAGC CTGCAGATGA GATCCGGGAG ACTGACACGA	1641
	AGCCAGCAGA GGTCAGAGGG GACTCTGACC ACAGCCCGCT CTCTGGCTGT CTGTCTGCAG	1701
35	TGCCCCGGCTG AGGGTGGGAG GCAAACACGC CTTGTTCTCG CTCTTCCCAG TTCAGCTTGG	1761
	TGGGAGAAAG TCATTCGCGT GGCTCGGGAC GCTCATGTAA ATTTGGTTTTT GGTGCTCAAG	1821
40	GGTTCCTTCC TCCCAGGGGC AGGTGTTTCT TTCCTGTTTG TCTTGTGTCT TGAGAGCTTG	1881
	GCCTTATGAC CAGTGAGAAC TCTCTCCCTG GTCTCTGCCA GCCCAAGCAT CACTGCCCCG	1941
	GGCGCCAGCT CAGTTTCACC GTCCACGTCC ACAAGGGGCT TTTCCACCT TCACCTTTGT	2001
45	CGCTGGGTCA GTGCTGGAAA GCGCCCTCA CTCCTGCGCT GACAAGGGCC CTTCTCTACT	2061
	GTCTGTGGGG TGGTTCCGGG CTGGGGGGGC TGCCTCCTTT GCACCTGATT TTGAAGGTGT	2121
50	CTCTTTCATC CATGGTTAAG TCATAAAAAG CTTATTGGTT TTGGTTTTGA CTCACCTGAA	2181
	AGTTTTTTTTG GTTTAAAAGA AGAATAGGCG GGGCACGGTG GCTCGTGCCT GTAATCCCAG	2241
	CACTTTGGGA GGCTGAGGCA GGTGGATCAC GAGGTCAGGA GATCGACACC ATCCTGGCTA	2301
55	ACACGGTGAA GCCCCGTCTC TACTAAAAAA TACAAAAAAT TAGCTGGGTG TGGTGGTGGG	2361
	GGTGGGCGCC TGTGGTCCCA GCTACGTGGG AGGCTGAGGC AGCAGACTGG TGTGAACCCG	2421
60	GGAGGTGGAG CTTGCAGTGA GCCGAGGTCG CGCCACTGCA CTCCAGCCTG GGCGACAGAG	2481
	CGAGACTCCA TCTCAAAAAA AAAAAA	2507

SUBSTITUTE SHEET (RULE 26)

## (2) INFORMATION FOR SEQ ID NO: 2:

## (i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 451 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## 10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Val Ser Arg Asp Gln Ala His Leu Gly Pro Lys Tyr Val Gly Leu  
1 5 10 15  
Trp Asp Phe Lys Ser Arg Thr Asp Glu Glu Leu Ser Phe Arg Ala Gly  
20 25 30  
Asp Val Phe His Val Ala Arg Lys Glu Glu Gln Trp Trp Trp Ala Thr  
35 40 45  
Leu Leu Asp Glu Ala Gly Gly Ala Val Ala Gln Gly Tyr Val Pro His  
50 55 60  
Asn Tyr Leu Ala Glu Arg Glu Thr Val Glu Ser Glu Pro Trp Phe Phe  
65 70 75 80  
Gly Cys Ile Ser Arg Ser Glu Ala Val Arg Arg Leu Gln Ala Glu Gly  
85 90 95  
Asn Ala Thr Gly Ala Phe Leu Ile Arg Val Ser Glu Lys Pro Ser Ala  
100 105 110  
Asp Tyr Val Leu Ser Val Arg Asp Thr Gln Ala Val Arg His Tyr Lys  
115 120 125  
Ile Trp Arg Arg Ala Gly Gly Arg Leu His Leu Asn Glu Ala Val Ser  
130 135 140  
Phe Leu Ser Leu Pro Glu Leu Val Asn Tyr His Arg Ala Gln Ser Leu  
145 150 155 160  
Ser His Gly Leu Arg Leu Ala Ala Pro Cys Arg Lys His Glu Pro Glu  
165 170 175  
Pro Leu Pro His Trp Asp Asp Trp Glu Arg Pro Arg Glu Glu Phe Thr  
180 185 190  
Leu Cys Arg Lys Leu Gly Ser Gly Tyr Phe Gly Glu Val Phe Glu Gly  
195 200 205  
Leu Trp Lys Asp Arg Val Gln Val Ala Ile Lys Val Ile Ser Arg Asp  
210 215 220  
Asn Leu Leu His Gln Gln Met Leu Gln Ser Glu Ile Gln Ala Met Lys  
225 230 235 240  
Lys Leu Arg His Lys His Ile Leu Ala Leu Tyr Ala Val Val Ser Val  
245 250 255  
Gly Asp Pro Val Tyr Ile Ile Thr Glu Leu Met Ala Lys Gly Ser Leu  
260 265 270  
Leu Glu Leu Leu Arg Asp Ser Asp Glu Lys Val Leu Pro Val Ser Glu  
275 280 285  
Leu Leu Asp Ile Ala Trp Gln Val Ala Glu Gly Met Cys Tyr Leu Glu  
290 295 300

33

Ser Gln Asn Tyr Ile His Arg Asp Leu Ala Ala Arg Asn Ile Leu Val  
 305 310 315 320  
 Gly Glu Asn Thr Leu Cys Lys Val Gly Asp Phe Gly Leu Ala Arg Leu  
 5 325 330 335  
 Ile Lys Glu Asp Val Tyr Leu Ser His Asp His Asn Ile Pro Tyr Lys  
 340 345 350  
 Trp Thr Ala Pro Glu Ala Leu Ser Arg Gly His Tyr Ser Thr Lys Ser  
 10 355 360 365  
 Asp Val Trp Ser Phe Gly Ile Leu Leu His Glu Met Phe Ser Arg Gly  
 15 370 375 380  
 Gln Val Pro Tyr Pro Gly Met Ser Asn His Glu Ala Phe Leu Arg Val  
 385 390 395 400  
 Asp Ala Gly Tyr Arg Met Pro Cys Pro Leu Glu Cys Pro Pro Ser Val  
 20 405 410 415  
 His Lys Leu Met Leu Thr Cys Trp Cys Arg Asp Pro Glu Gln Arg Pro  
 420 425 430  
 Cys Phe Lys Ala Leu Arg Glu Arg Leu Ser Ser Phe Thr Ser Tyr Glu  
 25 435 440 445  
 Asn Pro Thr  
 30 450

## (2) INFORMATION FOR SEQ ID NO: 3:

- 35 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 32 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 40 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

45 GGAATTCTAG AMGSGACYTG GCVGCBAGRA AC

32

## (2) INFORMATION FOR SEQ ID NO: 4:

- 50 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 32 base pairs  
 (B) TYPE: nucleic acid  
 55 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

60 GGAATTCTAG ACACSCCRWA RSWCCASACR TC

32

65

34

## (2) INFORMATION FOR SEQ ID NO: 5:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 8 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

His Arg Asp Leu Ala Ala Arg Asn  
1 5

## (2) INFORMATION FOR SEQ ID NO: 6:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 7 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(ix) FEATURE:

- (A) NAME/KEY: Modified-site  
(B) LOCATION: 5  
(D) OTHER INFORMATION: /note= "Xaa at position 5 is Tyr or Phe"

(ix) FEATURE:

- (A) NAME/KEY: Modified-site  
(B) LOCATION: 7  
(D) OTHER INFORMATION: /note= "Xaa at position 7 is Val or Ile"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

Asp Val Trp Ser Xaa Gly Xaa  
1 5

## (2) INFORMATION FOR SEQ ID NO: 7:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 35 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

GACTCGAGTC GACATCGATT TTTTTTTTTT TTTT

35

35

## (2) INFORMATION FOR SEQ ID NO: 8:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 19 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

CCTCGTCGGG GAAAACACC

19

## (2) INFORMATION FOR SEQ ID NO: 9:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

GGCTTATCAA GGAGGACGTC

20

## (2) INFORMATION FOR SEQ ID NO: 10:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 17 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

CTGGCCACGT GGAAGAC

17

## (2) INFORMATION FOR SEQ ID NO: 11:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 17 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

ACTTGAAGTC CCAGAGG

17

SUBSTITUTE SHEET (RULE 26)

36

## (2) INFORMATION FOR SEQ ID NO: 12:

## (i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 57 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## 10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

Ala Gly His Ala Gly Cys Ala Ala Leu Gln Asp Leu Ala Ala Cys Arg  
1 5 10 15  
Gly Pro Ala Ala Pro Glu Arg Gly Gly Val Leu Pro Gln Pro Ala Arg  
20 25 30  
Ala Cys Glu Leu Pro Gln Gly Pro Glu Pro Val Pro Arg Pro Ala Ala  
35 40 45  
20 Gly Arg Ala Leu Pro Glu Ala Arg Ala  
50 55

CLAIMS:

1. A DNA isolate encoding a protein tyrosine kinase or a fragment thereof, the protein tyrosine kinase being  
5 characterised by the amino acid sequence shown in  
SEQ ID NO. 2  
or by an amino acid sequence showing a significant degree of  
homology thereto.
- 10 2. A DNA isolate as claimed in Claim 1 wherein the amino  
acid sequence shows at least 60% homology in the catalytic  
domain and at least 40% homology in the remainder of the  
molecule to the sequence shown in SEQ ID NO. 2.
- 15 3. A DNA isolate as claimed in Claim 1 wherein the amino  
acid sequence shows at least 90% homology to the sequence  
shown in SEQ ID NO. 2.
- 20 4. A DNA isolate as claimed in any of Claims 1 to 3 encoding  
the full amino acid sequence of the protein tyrosine kinase.
- 25 5. A DNA isolate as claimed in any of Claims 1 to 3 encoding  
all or part of the catalytic domain of the protein tyrosine  
kinase.
- 30 6. A DNA isolate as claimed in any of claims 1 to 4 wherein  
the DNA encoding the protein tyrosine kinase or a fragment  
thereof has the DNA sequence shown in SEQ ID NO. 1.
- 35 7. A DNA isolate as claimed in any of claims 1 to 3 encoding  
amino acids 1 to 77 of  
SEQ ID NO. 2  
followed by the amino acids of  
SEQ ID NO. 12  
or an amino acid sequence showing a significant degree of  
homology thereto.
8. A recombinant expression vector comprising a DNA as

claimed in any of Claims 1 to 7 under control of an appropriate promoter and regulatory elements for expression in a host cell.

- 5 9. A recombinant cell line transformed with an expression vector as claimed in Claim 8, and capable of expressing a recombinant protein tyrosine kinase or fragment thereof characterised by the amino acid sequence shown in  
SEQ ID NO. 2,  
10 or an amino acid sequence showing a significant degree of homology thereto.
10. A recombinant cell line as claimed in Claim 9 which is a CHO, myeloma, primary immortalised breast, rodent  
15 fibroblast, baculovirus, yeast or *E. coli* cell line.
11. A protein tyrosine kinase or a fragment thereof characterised by the amino acid sequence shown in  
SEQ ID NO. 2  
20 or an amino acid sequence showing a significant degree of homology thereto.
12. A protein tyrosine kinase or a fragment thereof as claimed in Claim 11 which is a recombinant polypeptide.  
25
13. A protein tyrosine kinase or a fragment thereof as claimed in Claim 11 which has been produced in a human tumour cell line.
- 30 14. A protein tyrosine kinase or a fragment thereof having the amino acid sequence of amino acids 1 to 77 of  
SEQ ID NO. 2  
followed by the amino acids of  
SEQ ID NO. 12  
35 or an amino acid sequence showing a significant degree of homology thereto.
15. A method of screening a substance for potential utility



as a therapeutic agent in the treatment of cancer which comprises providing a standard system in which a protein tyrosine kinase or an active fragment thereof is able to develop a measurable effect, allowing the protein tyrosine kinase to develop that effect in the presence and absence of the said substance and measuring that effect, ability to produce significant inhibition of the effect being taken as an indication of potential utility as a therapeutic agent, wherein the protein tyrosine kinase is characterised by the amino acid sequence shown in

~~SEQ ID NO. 2~~

or an amino acid sequence showing a significant degree of homology thereto.

16. A method as claimed in Claim 15 wherein the protein tyrosine kinase or an active fragment thereof is derived from a recombinant expression system or a cellular source in which the protein is over-expressed.

17. A method as claimed in Claim 16 wherein the recombinant expression system is an insect baculovirus, an immortalized human breast cell line or a rodent fibroblast cell line.

18. A method as claimed in any of Claims 15 to 17 wherein the measurable effect is protein tyrosine kinase activity which is measured by measuring autophosphorylation or phosphorylation of a model substrate, by following incorporation of radioactive phosphate or by using anti-phosphotyrosine antibodies.

19. A method as claimed in Claim 15 wherein the measurable effect is a phenotypic alteration conferred by expression of the protein tyrosine kinase or an active fragment thereof in a recombinant or non-recombinant cell line.

20. A method as claimed in Claim 15 wherein the measurable effect is the tumorigenic effect of a recombinant or non-recombinant cell line expressing the protein tyrosine kinase

SUBSTITUTE SHEET (RULE 26)

or an active fragment thereof when used as a xenograft in an animal.

21. A method as claimed in Claim 19 or 20 wherein the  
5 recombinant cell line is a primary immortalised human breast cell line or a rodent fibroblast cell line.

22. A therapeutic agent identified by a method as claimed in  
any of Claims 15 to 21.

10

23. A therapeutic agent as claimed in Claim 22 which is a chemical molecule of relatively low molecular weight.

24. A therapeutic agent as claimed in Claim 23 which is a  
15 staurosporine analogue, a tyrphostin or a flavenoid.

25. A therapeutic agent as claimed in Claim 22 which is a peptide.

20 26. A therapeutic agent as claimed in Claim 22 which is an antisense oligonucleotide capable of blocking translation of the mRNA encoding the protein tyrosine kinase.

25 27. A method for detecting a tyrosine kinase having an amino acid sequence as defined in  
SEQ ID NO. 2

or a fragment thereof which comprises contacting a test sample with a specific antibody raised against an antigen from the said amino acid sequence and determining whether there is any  
30 antigen antibody binding within the test sample.

28. A method for the detection in a sample of a nucleic acid sequence encoding a tyrosine kinase which comprises subjecting the sample or nucleic acid isolated therefrom to a method  
35 capable of detecting a nucleic acid sequence therein identical or substantially homologous to a defined nucleic acid sequence and identifying any sequence identical or substantially homologous to all or part of the DNA sequence defined in

SEQ ID NO. 1,

DNA sequence from another part(s) of the tyrosine kinase characterised by that sequence or mRNA which would be the transcription product of such DNA.

5

29. A method as claimed in Claim 28 wherein DNA is detected by Southern blotting and/or PCR using appropriate primers.

10 30. A method as claimed in Claim 28 wherein RNA is detected by Northern blotting, RNase protection studies and direct nucleic acid *in situ* hybridisation.

15 31. An enzyme substrate complex which comprises a protein tyrosine kinase characterised by the amino acid sequence shown in

SEQ ID NO. 2

or by an amino acid sequence showing a significant degree of homology thereto and a therapeutic agent capable of modulating the activity of the said protein tyrosine kinase.

20

32. A complex as claimed in Claim 30 wherein the therapeutic agent is a chemical molecule of relatively low molecular weight.

25 33. A complex as claimed in Claim 32 wherein the therapeutic agent is a staurosporine analogue, a tyrphostin or a flavenoid.

30 34. A complex as claimed in Claim 30 wherein the therapeutic agent is a macromolecule.

35 35. A complex as claimed in Claim 34 wherein the therapeutic agent is an antibody.

36. A complex as claimed in Claim 30 wherein the therapeutic agent as a peptide.

37. A complex as claimed in Claim 30 wherein the therapeutic

SUBSTITUTE SHEET (RULE 26)

agent is an antisense oligonucleotide capable of blocking translation of the mRNA encoding the protein tyrosine kinase.

1/5

FIG. 1

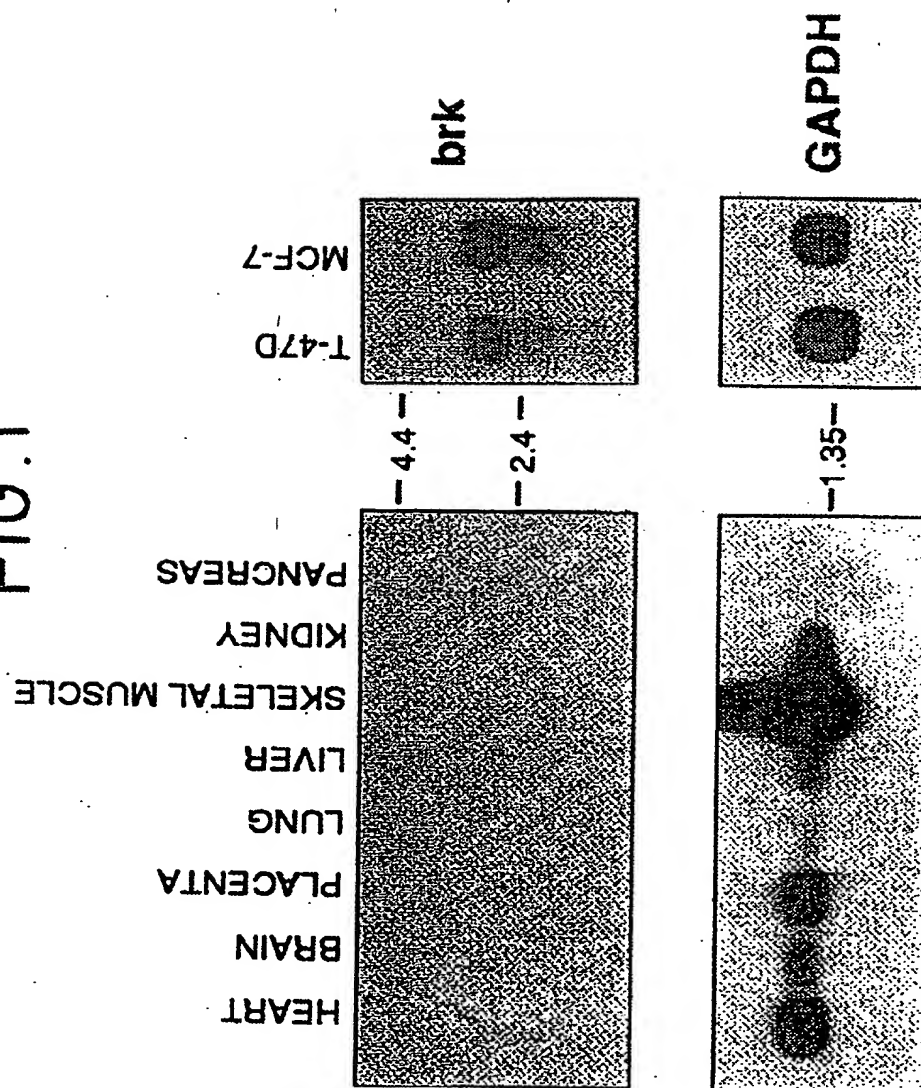
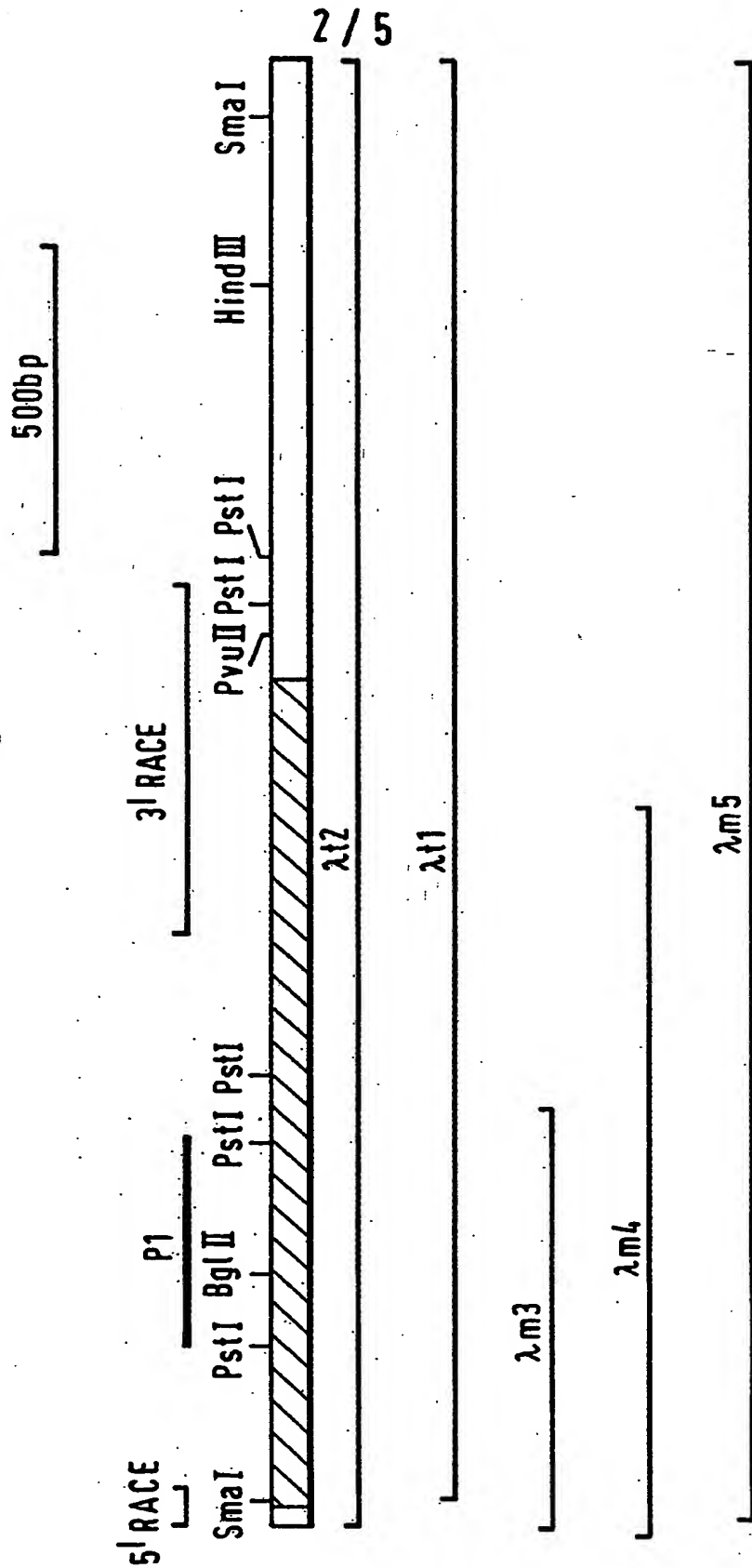
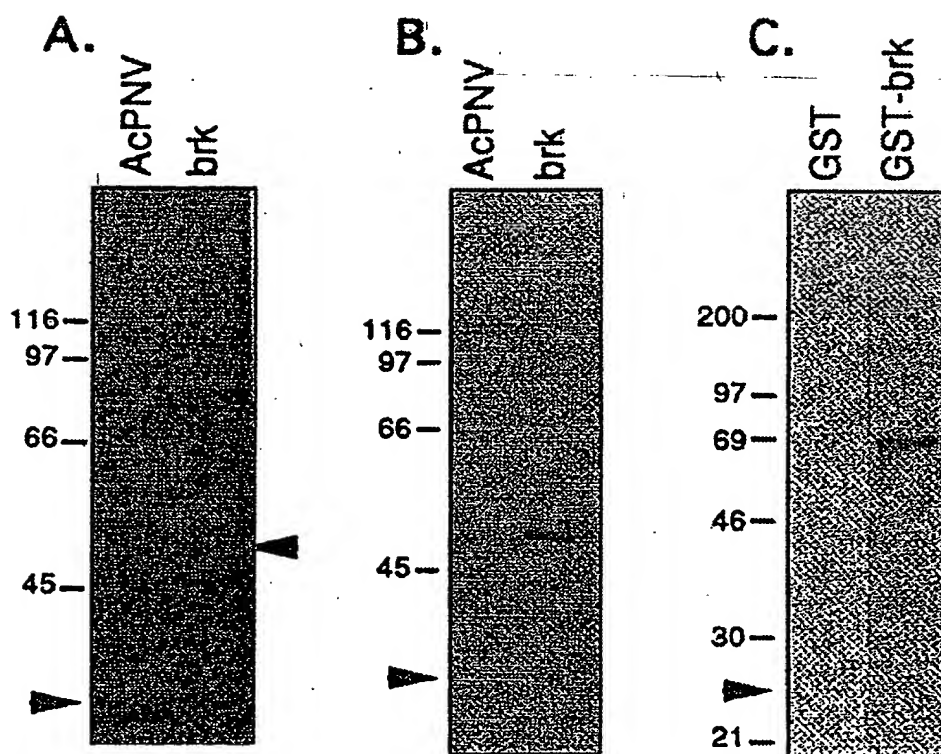


FIG. 2



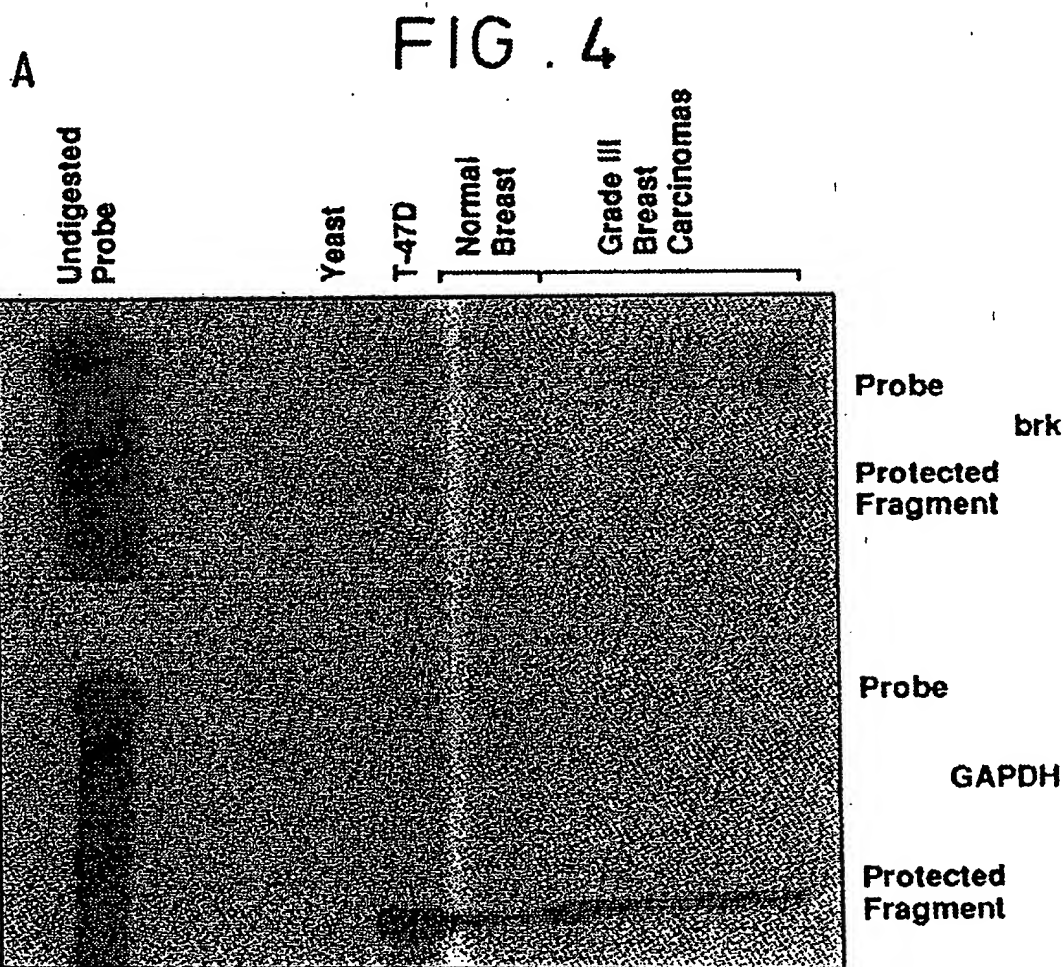
3 / 5

FIG. 3



SUBSTITUTE SHEET (RULE 26)

4/5





5/5

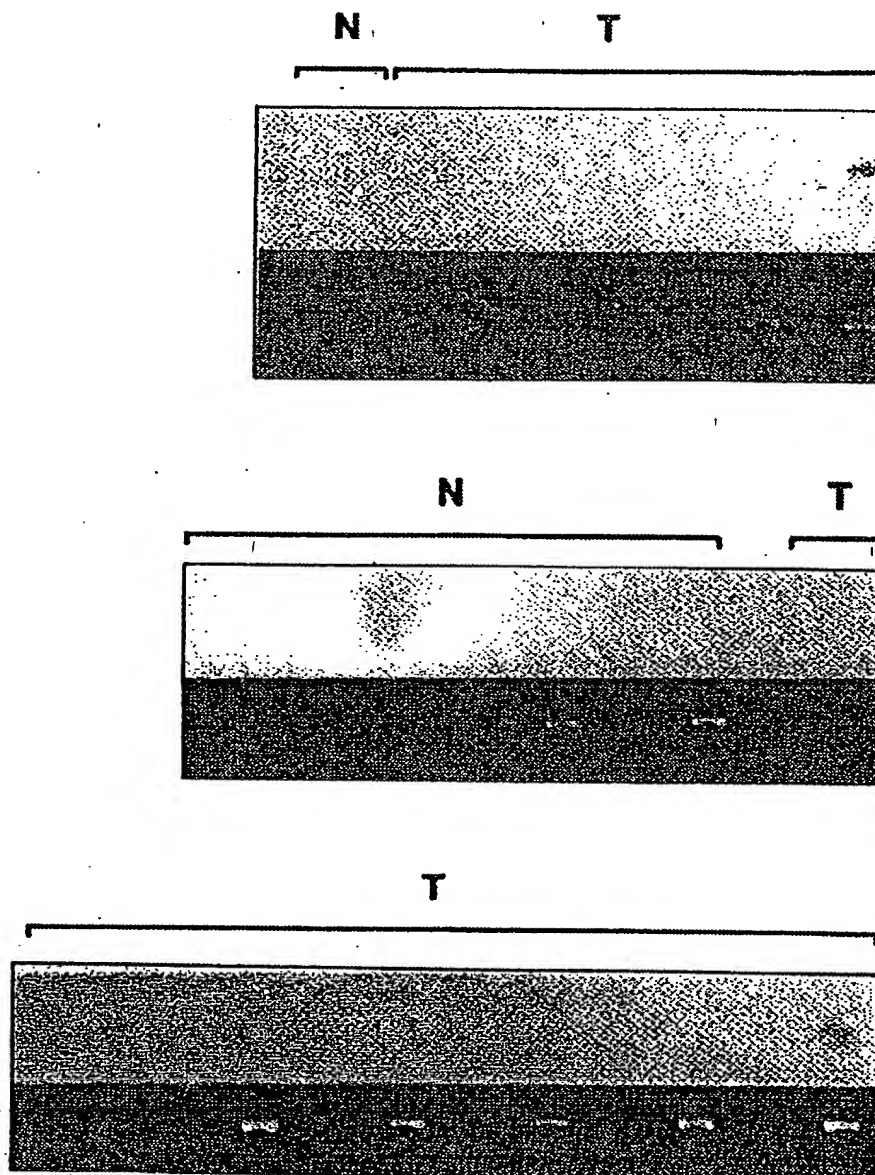


FIG. 5

SUBSTITUTE SHEET (RULE 26)

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/GB 94/01479

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 6	C12N15/54	C12N15/79	C12N9/12	C12N15/10	C12N1/21
	C12N1/19	C12Q1/48	C12Q1/68	G01N33/573	A61K38/17
	A61K31/70	A61K31/55	A61K31/275	A61K31/35	A61K47/48

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12N A61K G01N C12Q C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EP,A,0 296 110 (CIBA-GEIGI), 21 December 1988 see page 2 ---	22-24
X	J. BIOL. CHEM., vol.268, no.15, 25 May 1993 pages 11134 - 11142 OSHEROV ET AL. 'Selective inhibition of the epidermal growth factor and HER2/Neu receptors by tyrphostins' ---	22-24
A	---	27, 31-33,35
X	ANNUAL REPORTS IN MEDICINAL CHEMISTRY, vol.27, 1992 pages 169 - 178 DOBRUSIN ET AL. 'Chapter 18 : Protein tyrosine kinases and cancer' ---	22-25
	--- -/--	

☒ Further documents are listed in the continuation of box C.☒ Patent family members are listed in annex.

## \* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search

9 November 1994

Date of mailing of the international search report

29 -11- 1994

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Gac, G

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/GB 94/01479

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
T	ONCOGENE, vol.9, no.8, August 1994 pages 2383 - 2390 MITCHELL ET AL. 'Cloning and characterization of cDNAs encoding a novel non-receptor tyrosine kinase, brk, expressed in humn breast tumors' ----	1-6, 8-13, 28, 29
A	CA,A,2 054 602 (POWSON A.) 1 May 1993  see the whole document ----	15-20, 31, 35, 36
A	INT. J. CANCER, vol.54, no.4, 19 June 1993 pages 571 - 577 CANCE ET AL. 'Novel protein kinases expressed in human breast cancer' see the whole document ----	1-6, 8-13, 27-29
A	NATURE, vol.297, 20 May 1982 pages 205 - 208 KITAMURA ET AL. 'Avian sarcoma virus Y73 genome sequence and structural similarity of its transforming gene product to that of Rous sarcoma virus' see the whole document ----	1-6, 11-13
A	J. BIOL. CHEM., vol.267, no.7, 5 March 1992 pages 4357 - 4363 JALLAL ET AL. 'Tyrosine phosphatase inhibition permits analysis of signal transduction complexes in p185 HER2/neu-overexpressing human tumor cells' see the whole document ----	15, 16, 18, 20, 21, 27, 31, 32
A	WO,A,91 02085 (HOUSEY) 21 February 1991 see abstract see page 13, line 17 - line 26 ----	15, 19
A	WO,A,91 03260 (TEMPLE UNIVERSITY OF THE COMMON WEALTH SYSTEM OF HIGHER EDUCATION) 21 March 1991 see abstract ----	22, 26
A	EUR. J. BIOCHEM., vol.213, no.2, 15 April 1993 pages 701 - 710 DOROW ET AL. 'Identification of a new family of human epithelial protein kinases containing two Leucine/Isoleucone zipper domains' see the whole document -----	28-30

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.  
PCT/GB 94/01479

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP-A-0296110	21-12-88	AU-B- 617324 AU-A- 1757188 FI-B- 89362 JP-A- 1034989 US-A- 5093330	28-11-91 15-12-88 15-06-93 06-02-89 03-03-92
CA-A-2054602	01-05-93	NONE	
WO-A-9102085	21-02-91	US-A- 5266464 AU-A- 6427190 EP-A- 0486619 JP-T- 5503419	30-11-93 11-03-91 27-05-92 10-06-93
WO-A-9103260	21-03-91	AU-B- 646643 AU-A- 6410790 CA-A- 2065294 EP-A- 0489846 JP-T- 5500217	03-03-94 08-04-91 02-03-91 17-06-92 21-01-93